

Spacetime perspective of Schwarzschild lensing

To appear in *Physical Review D*

Simonetta Frittelli^{a,b}, Thomas P. Kling^b and Ezra T. Newman^b

^aDepartment of Physics, Duquesne University, Pittsburgh, PA 15282

^bDepartment of Physics and Astronomy, University of Pittsburgh, Pittsburgh, PA 15260
(November 29, 1999)

We propose a definition of an exact lens equation *without reference to a background spacetime*, and construct the exact lens equation explicitly in the case of Schwarzschild spacetime. For the Schwarzschild case, we give exact expressions for the angular-diameter distance to the sources as well as for the magnification factor and time of arrival of the images. We compare the exact lens equation with the standard lens equation, derived under the thin-lens-weak-field assumption (where the light rays are geodesics of the background with sharp bending in the lens plane, and the gravitational field is weak), and verify the fact that the standard weak-field thin-lens equation is inadequate at small impact parameter. We show that the second-order correction to the weak-field thin-lens equation is inaccurate as well. Finally, we compare the exact lens equation with the recently proposed strong-field thin-lens equation, obtained under the assumption of straight paths but without the small angle approximation, i.e., with allowed large bending angles. We show that the strong-field thin-lens equation is remarkably accurate, even for lightrays that take several turns around the lens before reaching the observer.

I. INTRODUCTION

The phenomenon of gravitational lensing is firmly associated with the physics of a four-dimensional Lorentzian spacetime that satisfies the Einstein equations. Yet, it has become a common practice in the study of lensing to break with the basic ideas of general relativity by using the linearized Einstein equations off a fixed background, the thin lens approximation, and treating the bending of light as a linear phenomenon – without mention of its connection with the full theory.

This point of view is very much justified by the accuracy in the comparison of contemporary observations with the resulting calculations, i.e., general relativity does play an essential role in lensing but the weak field approach appears to be quite adequate for most discussions [1].

However, it is now a fact that the strong field characteristics of general relativity *per se* are observed in nature as well. Black holes are possibly ubiquitous [2], and a super-massive black hole may exist in the center of every spiral galaxy. Here is where the full theory of general relativity takes the leading part. In order to describe bending of light by black holes or in high curvature regions, it is necessary to write lens equations that respect the intrinsic nature of general relativity, namely: covariance and non-linearity.

The difficulty in writing down a lens equation that respects covariance and non-linearity is very much of a conceptual type. In fact, even when such a lens equation is developed, it is hard to interpret. A spacetime containing a lens *is not* the superposition of two spaces, a background spacetime and a lens space. Two different spacetimes are two different entities, and there are an infinite number of ways of identifying them point-wise. What is the meaning of the angular location of a source *in the absence of a lens* – an idea used extensively in the thin-lens approximation? What are the preferred angular coordinates that give the thin-lens equation its meaning? How do we refer to the distances between the observer, the source and the lens in a coordinate independent manner, or what is the preferred coordinate distance to use? All these questions have perfectly good answers if a background spacetime is available to us and we are given leave to isolate the lensing action from the background. This is not so if there is no background. Without reference to a background, some of these questions have no answers, and some do not even make sense. Treating lensing phenomena strictly in the context of the full theory of relativity requires other ideas and approaches.

We have recently introduced a proposal for a lens equation without reference to a background [3,4]. An exact lens equation on an arbitrary Lorentzian spacetime can be written down, at least in principle, since it amounts, basically, to finding all the light-rays that reach the eye of an observer. However, for it to be meaningful, it is necessary to express the equation in such a way that it can be used in an astrophysical context; it must be written or expressed in terms of observable quantities. To some extent, we believe that we have partially succeeded in doing that.

As an illustration, we develop and interpret in full detail our lens equation in the case of a Schwarzschild black hole; explicitly working out quantities of astrophysical interest for lensing, such as the angular diameter distance and magnification factors. Furthermore, we use our exact lens equation to test the effectiveness of other lens equations

that can be written down in the case that a background is available, most notably the lens equation obtained recently by K. S. Virbhadra and G. F. R. Ellis [5].

In Sec.II we discuss the idealized situation where, in principle, the null geodesic equations can be solved exactly for a static metric and stationary source and show, again in principle, how a set of lens equations can be constructed, while in Sec.III these ideas are then applied to the Schwarzschild black hole lensing problem. In this section, the important physical quantities such as the angular-diameter distance to a source and the magnification factor are explicitly calculated. In the subsequent sections, we compare the exact results with the thin lens calculations.

II. THE EXACT LENS EQUATION

We begin with a four-dimensional static spacetime $(\mathfrak{M}, g_{ab}(x^a))$ with local coordinates x^a and consider an observer, at rest in the local coordinates, on a world-line given parametrically by $x_0^a(\tau)$, τ being the observer's proper-time. The observer, looking out, sees null geodesics reaching him from all past null directions, l^a . These observed directions, labeled by the spatial projections (orthogonal to the observer's velocity vector, $v^a = \frac{d}{d\tau}x_0^a(\tau)$) of the null vectors, can be taken as the two angular coordinates of the observer's (past) celestial sphere, (α_1, α_2) . The null geodesics of the past lightcones from the observer's worldline thus carry these labels; the points on each null geodesic are further labeled by the parameter along the curve, which we take to be an affine parameter s suitably normalized so that $l_a v^a = 1$. Thus the past lightcone of the observer has the form

$$x^a = X^a(x_0^a(\tau), \alpha_1, \alpha_2, s) \quad (1)$$

where the functions $X^a(x_0^a(\tau), \alpha_1, \alpha_2, s)$ satisfy the geodesic equation

$$\dot{X}^a \nabla_a \dot{X}^b = 0$$

with

$$\dot{X}^a = \frac{\partial}{\partial s} X^a(x_0^a(\tau), \alpha_1, \alpha_2, s)$$

and the null condition

$$g_{ab} \dot{X}^a \dot{X}^b = 0. \quad (2)$$

The local coordinates x^a can be chosen so that one of them, say x^0 , is timelike and the remaining three x^i are spacelike, $i = 1, 2, 3$. In this case, the function \dot{X}^0 does not vanish at any point. Although we are interested in the past lightcone, it is more straightforward to work in terms of the future lightcone. The direction in which the lightrays are traced is not important in the case of interest, namely, static spacetimes. Therefore, \dot{X}^0 is everywhere positive.

This means that x^0 increases monotonically with the affine parameter s . Because \dot{X}^0 does not vanish anywhere, then, by (2), at all points on a geodesic one of \dot{X}^i is non-zero (different i possibly in different sections of the geodesic). For definiteness, we label this spatial coordinate by $i = 1$. From the implicit function theorem, we have that

$$x^1 = X^1(x_0^a(\tau), \alpha_1, \alpha_2, s) \quad (3)$$

can be inverted to obtain

$$s = S(x_0(\tau), \alpha_1, \alpha_2, x^1). \quad (4)$$

The inversion will only be possible in patches, since it is possible that \dot{X}^1 vanishes at isolated points. This means that s will be, in general, a multiple-valued function of x^1 . Still, this inversion allows us to reparametrize the geodesics in terms of just our coordinates and observation angles:

$$x^0 = X^0(\tau, \alpha_1, \alpha_2, S(\tau, \alpha_1, \alpha_2, x^1)) \equiv \hat{X}^0(\tau, \alpha_1, \alpha_2, x^1), \quad (5)$$

$$x^1 = x^1 \quad (6)$$

$$x^A = X^A(\tau, \alpha_1, \alpha_2, S(\tau, \alpha_1, \alpha_2, x^1)) \equiv \hat{X}^A(\tau, \alpha_1, \alpha_2, x^1). \quad (7)$$

with $A = 2, 3$. The idea is now to treat these equations as if they determine a source at the spacetime point (x^0, x^1, x^A) in terms of the observable quantities $(\tau, \alpha_1, \alpha_2)$ where we have assumed, *for the moment*, that the coordinate value

for x^1 can be determined from observation. We will treat the source as slow moving or effectively at rest. In this case, Eq. (7) is defined as the lens equation.

More specifically, we interpret this lens equation as follows. Consider a source at a spatial location x^i , emitting light at time x^0 . We can think of the coordinate x^1 as a type of radial coordinate. The remaining two coordinates x^A are thus a type of angular coordinates. The emitted light arrives at the observer at a time τ , in a direction (α_1, α_2) . Eq. (7) expresses the angular location of a source at radial distance x^1 in terms of the observation angles (α_1, α_2) . On the other hand, Eq. (5) is an exact “time of arrival” equation; it relates the time of emission, x^0 , at the radial location, x^1 , with the observer’s proper time τ , and arrival direction, (α_1, α_2) .

The lens equation, Eq. (7), represents a map from the image (or observation) angles (α_1, α_2) to the source position angles, x^A . The map breaks down at locations where the determinant

$$J(\tau, \alpha_1, \alpha_2, x^1) \equiv \det \frac{\partial(x^2, x^3)}{\partial(\alpha_1, \alpha_2)} \quad (8)$$

vanishes. $J(\tau, \alpha_1, \alpha_2, x^1) = 0$ defines the caustics (a three surface in four-space) of the family of past lightcones of the observer. The direct consequence of the break-down of the map is that multiple images of the same source can be observed. More specifically, often one can see an image, in direction (α_1, α_2) , of an object that lies on a null geodesic before it reaches a caustic (in affine distance), while simultaneously seeing a different image, in a different direction (α'_1, α'_2) , from the same object along a different null geodesic, but, in this case, the object lies beyond the caustic in affine distance. The parity of the images is given by the sign of J . In background dependent calculations, J^{-1} is often interpreted as a magnification factor with respect to the “unlensed” source, but, as we have no background, this would not be appropriate here. (Note that $J(\tau, \alpha_1, \alpha_2, x^1)$ could have been calculated holding s fixed instead of fixed x^1 ; the vanishing of J is independent of that choice. This follows from the general theory of Lagrangian submanifolds and maps.)

The lens equation, Eq. (7), is not yet entirely usable since it involves the (up to now) unobservable quantity x^1 . However, x^1 can be expressed in terms of observable quantities through the use of the idea of distance. Though there are many definitions of distance in use in general relativity and astrophysics, several of them can be considered to be observable and we thus explore the feasibility of inferring x^1 from the considerations of distance. We will investigate a definition of distance which is observable, namely the so-called angular-diameter distance - there being several closely related distance definitions [4,1].

Since we have, in principle, exact expressions for the past lightcone of the observer in terms of parameters adapted to the null geodesic congruence, we have a natural way of expressing the angular-diameter distance to the source in exact form. The angular-diameter distance is defined [1] in terms of the infinitesimal area spanned by the observer’s geodesic congruence at the location of the source per infinitesimal solid angle at the observer’s location, namely:

$$D_A = \left| \frac{dA_s}{d\Omega_0} \right|^{\frac{1}{2}} \quad (9)$$

In order to calculate the area dA_s , we define two connecting vectors in the lightcone of the observer. By taking variations of the points on the lightcone with respect to the labels of the null geodesics in the congruence, we find the geodesic deviation vectors, or Jacobi fields:

$$M_1^a = \frac{\partial X^a}{\partial \alpha_1}, \quad M_2^a = \frac{\partial X^a}{\partial \alpha_2}. \quad (10)$$

It is irrelevant to the area calculation whether s or x^1 are held constant in calculating the connecting vectors – the difference, lying along the null tangent vectors to the geodesics, does not affect the area.

The area dA_s is the area spanned by these two vectors at the location of the source, namely, the norm of the wedge-product of the two vectors:

$$\begin{aligned} dA_s &= \left| g_{ac} g_{bd} M_1^{[a} M_2^{b]} M_1^{[c} M_2^{d]} \right|^{\frac{1}{2}} d\alpha_1 d\alpha_2 \\ &= \left| 2 \left((M_1 \cdot M_1) (M_2 \cdot M_2) - (M_1 \cdot M_2)^2 \right) \right|^{\frac{1}{2}} d\alpha_1 d\alpha_2. \end{aligned} \quad (11)$$

If the solid angle at the observer subtended by the area dA_s is given by $d\Omega_0 = K(\alpha_1, \alpha_2) d\alpha_1 d\alpha_2$, where K depends on the choice of the coordinates, the angular-diameter distance is given by

$$D_A^2 = D_A^2(\alpha_1, \alpha_2, \tau, x_1) = 2 K^{-2} \left| (M_1 \cdot M_1) (M_2 \cdot M_2) - (M_1 \cdot M_2)^2 \right|. \quad (12)$$

For sufficiently small values of x^1 , Eq. (12) is invertible, i.e., $x^1 = x^1(\alpha_1, \alpha_2, \tau, D_A)$. However, D_A goes to zero at the caustic, so that beyond the caustic x^1 is a multivalued function of D_A and must be given in patches.

The angular-diameter distance is observable, because it is related to the intrinsic luminosity L of the source and its apparent brightness S (total flux at the observer) via [1]

$$S = \frac{L}{4\pi(1+z)^4 D_A^2}. \quad (13)$$

In principle, Eq. (13), with Eq. (12), gives x^1 implicitly as a function of observables: the angular location of the image (α_1, α_2) , its redshift $z(= \omega_s/\omega_0 - 1)$ its apparent brightness and the intrinsic luminosity of the source.

On the other hand, there may be situations where the intrinsic luminosity of the source is not available. In such cases, if there are multiple images observed, then we can make use of their relative brightness in order to estimate x^1 . For two images of the same source, lying at angles $(\alpha_1^{(1)}, \alpha_2^{(1)})$ and $(\alpha_1^{(2)}, \alpha_2^{(2)})$, the ratio of the fluxes S_1/S_2 does not depend on the intrinsic luminosity of the source L and can be interpreted as the relative magnification μ_{12} of one image with respect to the other one, or

$$\mu_{12} = \frac{D_A^2(\alpha_1^{(2)}, \alpha_2^{(2)}, x^1)}{D_A^2(\alpha_1^{(1)}, \alpha_2^{(1)}, x^1)}. \quad (14)$$

Notice that $(\alpha_1^{(1)}, \alpha_2^{(1)})$ and $(\alpha_1^{(2)}, \alpha_2^{(2)})$ are two image directions of Eq. (7) for a given value of the source coordinates (x^2, x^3) . The inversion of Eq. (14) is not likely to be feasible in closed form. Still, in principle, Eq. (14) gives x^1 implicitly in terms of the angular location of two images and their relative brightness $\mu_{12} \equiv S_1/S_2$.

(In the case of lensing at cosmological distances, it is customary to infer distances from redshifts. Even though we are not concerned with cosmological models in this paper, we consider redshifts as another alternative to infer x^1 from an observable quantity. For a source on a worldline with tangent vector v_s^a , emitted light of frequency ω_s and observed frequency ω_0 , the ratio ω_s/ω_0 is given by [6]

$$\frac{\omega_s}{\omega_0} = \frac{g_{ab}(x^a(\alpha_1, \alpha_2, x^1))v_s^a \dot{X}^b(\alpha_1, \alpha_2, x^1)}{g_{ab}(x_0^a)v_0^a \dot{X}^b(x_0^a)} \quad (15)$$

where Eq. (7) has been used for the source's space-time location. Equation (15) gives x^1 implicitly in terms of the frequency of emission, the received frequency and the observed image angle. As our assumed source is at rest, its velocity is $v_s^a = |g_{00}(x^a(\alpha_1, \alpha_2, x^1))|^{-1/2}(1, 0, 0, 0)$.

III. THE SCHWARZSCHILD CASE

A. The Lens Equation

We consider now the case of gravitational lensing by a Schwarzschild black hole of mass M . The line element is

$$ds^2 = f(r)dt^2 - \frac{1}{f(r)}dr^2 - r^2(d\theta^2 + \sin^2\theta d\phi^2) \quad (16)$$

with

$$f(r) \equiv 1 - \frac{2M}{r} \quad (17)$$

In order to take advantage of existing calculations [7], we temporarily use coordinates (u, l) given by

$$u = \frac{1}{\sqrt{2}} \left(t - \int \frac{dr}{f} \right) = \frac{1}{\sqrt{2}} (t - r + 2M \log(2M - r)), \quad (18)$$

$$l = \frac{1}{\sqrt{2}r}. \quad (19)$$

In these coordinates, the line element takes the form

$$ds^2 = 2f du^2 - \frac{2}{l^2} du dl - \frac{1}{2l^2} (d\theta^2 + \sin^2 \theta d\phi^2). \quad (20)$$

The equations for *null* geodesics $\ddot{x}^a + \Gamma_{bc}^a \dot{x}^b \dot{x}^c = 0$ in terms of an affine parameter s are equivalent to

$$\dot{u} = \frac{C}{2f} \left(1 \pm \sqrt{1 - \left(\frac{B}{C}\right)^2 l^2 f} \right) \quad (21)$$

$$\dot{l} = \pm Cl^2 \sqrt{1 - \left(\frac{B}{C}\right)^2 l^2 f} \quad (22)$$

$$\dot{\phi} = \frac{Al^2}{\sin^2 \theta} \quad (23)$$

$$\left(\frac{\dot{\theta}}{l^2}\right)^2 = B^2 - \frac{A^2}{\sin^2 \theta} \quad (24)$$

with the null condition $\dot{x}^a \dot{x}_a = 0$ equivalent to

$$4l^2 f \dot{u}^2 - 4Cl^2 \dot{u} + \dot{\theta}^2 + \frac{A^2 l^4}{\sin^2 \theta} = 0. \quad (25)$$

The symbol $(\dot{})$ stands for d/ds and A, B, C are three first integrals of the null geodesics, depending on the initial point and the initial direction. The constant C represents the freedom in the scaling of the affine parameter s . Treating the observers location $(u_0, l_0, \theta_0, \phi_0)$ as the initial point, the constant B is related to the angle ψ that the null geodesic makes with the optical axis (defined by the radial line from observer to the lens center). More precisely:

$$\frac{B}{C} = \frac{\sin \psi}{l_0 \sqrt{f(l_0)}}, \quad (26)$$

where l_0 is the inverse radial location of the observer. Lastly, the constant A can be related to the azimuthal angle γ that the direction of the lightray makes around the optical axis at the observer's location via

$$\frac{A}{C} = \sin \theta_0 \sin \gamma \frac{\sin \psi}{l_0 \sqrt{f(l_0)}}, \quad (27)$$

where θ_0 is the angular location of the observer.

We can switch from the null coordinate u to the time coordinate t using $t = \sqrt{2} \left(\dot{u} - \frac{1}{2l^2 f} \dot{l} \right)$. Doing so and setting $C = 1$ allows Eqs. (21-24) to be rewritten as

$$\dot{t} = \frac{1}{\sqrt{2}f} \quad (28)$$

$$\dot{l} = \pm l^2 \sqrt{1 - \left(\frac{\sin^2 \psi}{l_0^2 f(l_0)}\right) l^2 f} \quad (29)$$

$$\dot{\phi} = \frac{\sin \theta_0 \sin \gamma \sin \psi l^2}{l_0 \sqrt{f(l_0)} \sin^2 \theta} \quad (30)$$

$$\left(\frac{\dot{\theta}}{l^2}\right)^2 = \frac{\sin^2 \psi}{l_0^2 f(l_0)} - \frac{(\sin \theta_0 \sin \gamma \sin \psi)^2}{l_0^2 f(l_0) \sin^2 \theta} \quad (31)$$

For those null geodesics of interest to us, i.e., those whose initial direction has a component pointing towards the Schwarzschild origin, the inverse radial distance l initially increases ($\dot{l} > 0$) until the point of closest approach to the lens is reached, (with affine parameter value s_p). The coordinate, l , then decreases ($\dot{l} < 0$) after s_p until reaching the source at some s_{fin} . However, for $s < s_p$ and for $s > s_p$ we have $\dot{l} \neq 0$, and thus the inverse radial distance l can be used as a parameter (in two patches, the incoming and outgoing) along the null geodesics for the purposes of constructing our lens equation in the manner of the previous Section. It plays the role of x^1 .

The value of l at the point of closest approach (i.e. at $\dot{l} = 0$) is denoted l_p . If we assume, naturally, that the observer is located outside the last stable orbit (at $r > 3M$), then, for lightrays that do not cross the $r = 3M$ orbit, the closest approach l_p is the smallest of the positive roots of

$$(\sin \psi)^2 l_p^2 (1 - 2\sqrt{2}Ml_p) - l_0^2 (1 - 2\sqrt{2}Ml_0) = 0. \quad (32)$$

A simple analysis shows that for $l_0 < 3\sqrt{2}M$, Eq. (32) has no positive roots unless

$$\sin \psi > 3\sqrt{2}Ml_0 \sqrt{3(1 - 2\sqrt{2}Ml_0)} \quad (33)$$

in which case there are always two positive roots, and the closest approach l_p is the smallest of them. It is simple to prove that $l_p < (3\sqrt{2}M)^{-1}$ for all ψ subject to Eq. (33), and that $l_p \rightarrow (3\sqrt{2}M)^{-1}$ for $\sin \psi \rightarrow 3\sqrt{2}Ml_0 \sqrt{3(1 - 2\sqrt{2}Ml_0)}$. See Fig. 1.

The l_p , which is a turning point of the coordinate l along the null geodesics, plays a major role in the following. First we notice that the term,

$$1 - \left(\frac{\sin^2 \psi}{l_0^2 f(l_0)} \right) l^2 f$$

from the \dot{u} equation, can be rewritten with the role of ψ now played by l_p , in the form

$$1 - \left(\frac{\sin^2 \psi}{l_0^2 f(l_0)} \right) l^2 f = \frac{1}{l_p^2 (1 - 2\sqrt{2}Ml_p)} \left(l_p^2 (1 - 2\sqrt{2}Ml_p) - l^2 (1 - 2\sqrt{2}Ml) \right) \quad (34)$$

using, from Eq. (32),

$$\frac{\sin^2 \psi}{l_0^2 (1 - 2\sqrt{2}Ml_0)} = \frac{1}{l_p^2 (1 - 2\sqrt{2}Ml_p)}. \quad (35)$$

We thus see that the dependence on ψ is now hidden away in the inverse radial distance of closest approach l_p . This observation will simplify some of the calculations that follow.

The past lightcone of an observer in coordinates $x^a = (t, l, \theta, \phi)$ in terms of the affine parameter s and initial directions (ψ, γ) could in principle be obtained from Eq. (21 - 24) or Eq. (28 - 31). This requires the integration of four non-linear ordinary differential equations which can not be done by quadratures. In the spirit of the Section II, however, we do not need the lightcone in terms of the affine parameter s , but in terms of a radial coordinate. Our radial coordinate is the inverse radial distance l , which is better suited for treating large distances than the standard r . In particular, the infinite range $3M < r < \infty$ translates into the finite interval $0 < l < (3\sqrt{2}M)^{-1}$.

First we show how our radial coordinate l is related to the affine length. Next, we integrate the lightcone in terms of l . For our purposes, it suffices to assume that by the time the lightray reaches the observer it has already passed by the point of closest approach l_p .

Eq. (29) can be integrated to obtain the affine parameter s in terms of the inverse radial distance:

$$s = 2 \int_{l_0}^{l_p} \sqrt{\frac{l_p^2 (1 - 2\sqrt{2}Ml_p)}{l_p^2 (1 - 2\sqrt{2}Ml_p) - l'^2 (1 - 2\sqrt{2}Ml')}} \frac{dl'}{l'^2} + \int_l^{l_0} \sqrt{\frac{l_p^2 (1 - 2\sqrt{2}Ml_p)}{l_p^2 (1 - 2\sqrt{2}Ml_p) - l'^2 (1 - 2\sqrt{2}Ml')}} \frac{dl'}{l'^2}. \quad (36)$$

Equation (36) corresponds to Eq. (4) of the previous section. This represents one of the two available patches for s as a function of l .

The affine parameter as a function of l is represented in Fig. 2 when the observer is at a distance of $30M$. (We chose a relatively small distance in order to better appreciate the strong field effects.) The affine length goes to infinity as l approaches zero, in agreement with the fact that the lightray runs out to infinity. The affine length, s , is chosen to vanish at the observer's location. We see, in the diagram, that the affine length bulges towards the $3M$ radius, resulting in a double valued function of l . The bulge is more pronounced for lightrays that reach the observer at

smaller observation angles ψ . The rays that come closer to the $3M$ radius spend more affine time in reaching the observer, in agreement with the gravitational time delay.

In the following, we obtain the past lightcone of the observer $(t_0, l_0, \theta_0, \phi_0)$ as a function of the inverse radial distance l , instead of the affine parameter s , and two angles (ψ, γ) specifying the direction of each null geodesic at the observer's location.

Integrating Eq. (28) with Eq. (29), we obtain

$$t = t_0 + 2 \int_{l_0}^{l_p} \sqrt{\frac{l_p^2(1 - 2\sqrt{2}Ml_p)}{l_p^2(1 - 2\sqrt{2}Ml_p) - l^2(1 - 2\sqrt{2}Ml)}} \frac{dl'}{\sqrt{2}l'^2(1 - 2\sqrt{2}Ml')} \\ + \int_l^{l_0} \sqrt{\frac{l_p^2(1 - 2\sqrt{2}Ml_p)}{l_p^2(1 - 2\sqrt{2}Ml_p) - l'^2(1 - 2\sqrt{2}Ml')}} \frac{dl'}{\sqrt{2}l'^2(1 - 2\sqrt{2}Ml')} \quad (37)$$

This is the equivalent of Eq. (5) of the previous section. As a function of l , the time of arrival is double-valued (not so as a function of the affine parameter s); Eq. (37) represents one of the two patches.

The integration of the angular coordinates of the lightcone is carried out in [7]. Representing the angular coordinates (θ, ϕ) in terms of the complex stereographic variables $\zeta \equiv \cot(\theta/2)e^{i\phi}$ the integration yields

$$\zeta = e^{i\phi_0} \frac{\cot \frac{\theta_0}{2} + e^{i\gamma} \cot \frac{\Theta(l, l_0, l_p)}{2}}{1 - e^{i\gamma} \cot \frac{\Theta(l, l_0, l_p)}{2} \cot \frac{\theta_0}{2}} \quad (38)$$

where $\Theta(l, l_0, l_p)$ is

$$\Theta(l, l_0, l_p) = \pm \left(\pi - 2 \int_{l_0}^{l_p} \frac{dl}{\sqrt{l_p^2(1 - 2\sqrt{2}Ml_p) - l^2(1 - 2\sqrt{2}Ml)}} \right. \\ \left. - \int_l^{l_0} \frac{dl'}{\sqrt{l_p^2(1 - 2\sqrt{2}Ml_p) - l'^2(1 - 2\sqrt{2}Ml')}} \right). \quad (39)$$

The function $\Theta(l, l_0, l_p)$ depends on the observation angle ψ through l_p . The overall positive sign is taken when the value of the observation angle, ψ , is positive, and the negative sign is taken for negative ψ . This makes $\Theta(l, l_0, l_p)$ an odd function of ψ .

Geometrically, $\Theta(l, l_0, l_p)$ “represents” the angular position of the source relative to the optical axis, defined by the line between the lens and observer. The observer is considered to lie on the optical axis at $\Theta = \pi$. The relative angular position of a source is given by Θ values between $-\pi$ and π . Hence, Eq. (39) must be considered mod 2π , where values outside the range, $-\pi \leq \Theta \leq \pi$, represent multiple circlings of the lens. When $\Theta = 0, 2\pi, 4\pi, \dots$, the source is colinear with the lens and observer and would be observed as an Einstein Ring. For positive ψ , a value of $\Theta \bmod 2\pi$ between π and 0 represents a source located to the right of the lens, while $-\pi < \Theta \bmod 2\pi < 0$ represents a source located to the left of the optical axis.

Figure 3 shows a plot of Θ at fixed values of l and l_0 , as a function of the image angle ψ . We can see that Θ blows up at $l_p = (3\sqrt{2}M)^{-1}$, which agrees with the fact that, as the lightrays approach the $3M$ radius, they take a larger number of turns around the lens. Notice that Θ is a regular function of l for all $l < l_p$ because the integrand diverges slowly, as $(l_p - l)^{-1/2}$. In fact, for numerical integration it turns out to be much more efficient to make a change of variables $l = l_p - q$ and write Θ as

$$\Theta = \pm \left(\pi - 2 \int_0^{l_p - l_0} \frac{dq}{\sqrt{2l_p(1 - 2\sqrt{2}Ml_p)q + (6\sqrt{2}Ml_p - 1)q^2 - 2\sqrt{2}Mq^3}} \right. \\ \left. - \int_l^{l_0} \frac{dl'}{\sqrt{l_p^2(1 - 2\sqrt{2}Ml_p) - l'^2(1 - 2\sqrt{2}Ml')}} \right). \quad (40)$$

In terms of the standard spherical coordinates (θ, ϕ) , Eq. (38) translates into:

$$\begin{aligned}\cos \theta &= -\cos \theta_0 \cos \Theta + \sin \theta_0 \sin \Theta \cos \gamma \\ \tan \phi &= \frac{\sin \phi_0 \sin \theta_0 - \tan \Theta (\cos \phi_0 \sin \gamma - \sin \phi_0 \cos \gamma \cos \theta_0)}{\cos \phi_0 \sin \theta_0 + \tan \Theta (\sin \phi_0 \sin \gamma + \cos \phi_0 \cos \gamma \cos \theta_0)}\end{aligned}\quad (41)$$

Equations (41), with Eq. (40) are the exact lens equations for the case of a Schwarzschild spacetime and correspond to Eqs. (7) of the previous section. Notice that the observer is located at generic values of (θ_0, ϕ_0) , which means that we do not choose, as is often done, the z -axis as the optical axis, the optical axis being the radial line that contains both the center of symmetry and the observer. This is because the spherical coordinates break down along the z -axis. If we chose the observer to lie along the z -axis, then Eq. (41) reduce to $\cos \theta = \cos \Theta$ and $\tan \phi = \tan \gamma$, and we could interpret Θ and γ as the lens angular coordinates. However, this would result in erroneous predictions in the following subsections, unless we use additional care. In order to keep the remainder of this paper in the most transparent form, we prefer to keep the observer off the z -axis.

B. Lensing Observables

In this subsection, we describe the calculation of three key lensing observables from the lens equations: the angular-diameter distance, the relative magnifications, and the time-delay between the arrival times of two images.

We start by exploring the angular-diameter distance, using Eqs. (41) and Eq. (37) to obtain an exact expression of the angular-diameter distance in terms of the inverse parameter l . In the next subsection, we will use the expressions we obtain here to explore the possibility of inferring the inverse radial distance, l , to the source.

First, we define the connecting vectors

$$M_1^a \equiv \left(\frac{\partial t}{\partial \gamma}, \frac{\partial l}{\partial \gamma}, \frac{\partial \theta}{\partial \gamma}, \frac{\partial \phi}{\partial \gamma} \right) = \left(0, 0, \frac{\partial \theta}{\partial \gamma}, \frac{\partial \phi}{\partial \gamma} \right) \quad (42)$$

$$M_2^a \equiv \left(\frac{\partial t}{\partial \psi}, \frac{\partial l}{\partial \psi}, \frac{\partial \theta}{\partial \psi}, \frac{\partial \phi}{\partial \psi} \right) = \left(\frac{\partial t}{\partial \psi}, 0, \frac{\partial \theta}{\partial \psi}, \frac{\partial \phi}{\partial \psi} \right) \quad (43)$$

where t, θ, ϕ are functions of (l, ψ, γ) given by Eq. (37) and Eq. (41). The partial derivatives are taken at fixed value of l . From the expressions above, we have

$$\frac{\partial \theta}{\partial \gamma} = \frac{\sin \Theta \sin \theta_0 \sin \gamma}{\sqrt{1 - (\cos \Theta \cos \theta_0 - \sin \Theta \sin \theta_0 \cos \gamma)^2}} \quad (44)$$

$$\frac{\partial \theta}{\partial \psi} = -\frac{\sin \Theta \cos \theta_0 + \cos \Theta \sin \theta_0 \cos \gamma}{\sqrt{1 - (\cos \Theta \cos \theta_0 - \sin \Theta \sin \theta_0 \cos \gamma)^2}} \frac{\partial \Theta}{\partial \psi} \quad (45)$$

$$\frac{\partial \phi}{\partial \gamma} = -\frac{\sin \Theta (\sin \Theta \cos \theta_0 + \cos \Theta \sin \theta_0 \cos \gamma)}{1 - (\cos \Theta \cos \theta_0 - \sin \Theta \sin \theta_0 \cos \gamma)^2} \quad (46)$$

$$\frac{\partial \phi}{\partial \psi} = -\frac{\sin \theta_0 \sin \gamma}{1 - (\cos \Theta \cos \theta_0 - \sin \Theta \sin \theta_0 \cos \gamma)^2} \frac{\partial \Theta}{\partial \psi} \quad (47)$$

Notice that, by Eq. (44) and Eq. (46), the vector M_1^a is proportional to $\sin \Theta$ for all generic values of θ_0 except for $\theta_0 = 0, \pi$. This means that, generically, the vector M_1^a vanishes at $\Theta = 0$, which, by Eq. (41), represents source points (θ, ϕ) along the optical axis. If we had chosen the optical axis as the z -axis this essential fact would not be as transparent.

With the metric, Eq. (16), we have

$$M_1 \cdot M_1 = -\frac{1}{2l^2} \left(\left(\frac{\partial \theta}{\partial \gamma} \right)^2 + \sin^2 \theta \left(\frac{\partial \phi}{\partial \gamma} \right)^2 \right) \quad (48)$$

$$M_2 \cdot M_2 = f \left(\frac{\partial t}{\partial \psi} \right)^2 - \frac{1}{2l^2} \left(\left(\frac{\partial \theta}{\partial \psi} \right)^2 + \sin^2 \theta \left(\frac{\partial \phi}{\partial \psi} \right)^2 \right) \quad (49)$$

$$M_1 \cdot M_2 = -\frac{1}{2l^2} \left(\frac{\partial \theta}{\partial \gamma} \frac{\partial \theta}{\partial \psi} + \sin^2 \theta \frac{\partial \phi}{\partial \gamma} \frac{\partial \phi}{\partial \psi} \right). \quad (50)$$

Using Eqs. (48 - 50) the area dA_s from Eq. (11) can be written as

$$dA_s = \left(\frac{\sin^2 \theta}{4l^4} \left(\frac{\partial \theta}{\partial \gamma} \frac{\partial \phi}{\partial \psi} - \frac{\partial \phi}{\partial \gamma} \frac{\partial \theta}{\partial \psi} \right)^2 + f \left(\frac{\partial t}{\partial \psi} \right)^2 M_1 \cdot M_1 \right)^{\frac{1}{2}} d\psi d\gamma. \quad (51)$$

The determinant of the lens map J (Eq. (8)),

$$J = \frac{\partial \theta}{\partial \gamma} \frac{\partial \phi}{\partial \psi} - \frac{\partial \phi}{\partial \gamma} \frac{\partial \theta}{\partial \psi}$$

which appears in Eq. (51) can be simplified using Eqs. (44 - 47):

$$J = \frac{\sin \Theta}{\sqrt{1 - (\cos \Theta \cos \theta_0 - \sin \Theta \sin \theta_0 \cos \gamma)^2}} \frac{\partial \Theta}{\partial \psi}. \quad (52)$$

The scalar product $M_1 \cdot M_1$ can also be evaluated using Eqs. (44 - 47):

$$M_1 \cdot M_1 = -\frac{\sin^2 \Theta}{2l^2} \quad (53)$$

Thus Eq. (51) becomes

$$dA_s = \frac{\sin \Theta}{2l^2} \left(\left(\frac{\partial \Theta}{\partial \psi} \right)^2 - 2l^2 f \left(\frac{\partial t}{\partial \psi} \right)^2 \right)^{\frac{1}{2}} d\psi d\gamma \quad (54)$$

The solid angle at the observer's location is $d\Omega_0 = \sin \psi d\psi d\gamma$. The angular-diameter distance is thus

$$D_A^2 = \frac{\sin \Theta}{2l^2 \sin \psi} \left(\left(\frac{\partial \Theta}{\partial \psi} \right)^2 - 2l^2 f \left(\frac{\partial t}{\partial \psi} \right)^2 \right)^{\frac{1}{2}} \quad (55)$$

This expression for the angular-diameter distance can be simplified by showing that $\partial t / \partial \psi$ can be expressed as a linear function of $\partial \Theta / \partial \psi$. We present a short, intuitive derivation of this fact here; a more formal derivation is given in the appendix.

First, we notice that both connecting vectors M_1^a and M_2^a lie on the lightcone and therefore must be orthogonal to the null vector that is tangent to the lightrays, with components given by $\ell^a \equiv (\dot{t}, \dot{l}, \dot{\theta}, \dot{\phi})$. The scalar product of M_2^a with ℓ^a is

$$g_{ab} M_2^a \ell^b = f \dot{t} \frac{\partial t}{\partial \psi} - \frac{1}{2l^2} \left(\dot{\theta} \frac{\partial \theta}{\partial \psi} + \sin^2 \theta \dot{\phi} \frac{\partial \phi}{\partial \psi} \right) \quad (56)$$

where $\dot{\theta}$ and $\dot{\phi}$ are explicitly given by

$$\begin{aligned} \dot{\theta} &= \frac{\sin \Theta \cos \theta_0 + \cos \Theta \sin \theta_0 \cos \gamma}{\sqrt{1 - (\cos \Theta \cos \theta_0 - \sin \Theta \sin \theta_0 \cos \gamma)^2}} \frac{\sin \psi}{\sqrt{l_0^2 (1 - 2\sqrt{2} M l_0)}} \\ \dot{\phi} &= \frac{\sin \theta_0 \sin \gamma}{1 - (\cos \Theta \cos \theta_0 - \sin \Theta \sin \theta_0 \cos \gamma)^2} \frac{\sin \psi}{\sqrt{l_0^2 (1 - 2\sqrt{2} M l_0)}} \end{aligned} \quad (57)$$

By inserting Eq. (45), Eq. (47), and Eqs. (57) into Eq. (56), we have

$$g_{ab} M_2^a \ell^b = \frac{1}{\sqrt{2}} \frac{\partial t}{\partial \psi} + \frac{\sin \psi}{2\sqrt{l_0^2 (1 - 2\sqrt{2} M l_0)}} \frac{\partial \Theta}{\partial \psi}. \quad (58)$$

Since $g_{ab} M_2^a \ell^b = 0$, we obtain the claimed result

$$\frac{\partial t}{\partial \psi} = -\frac{\sin \psi}{\sqrt{2l_0^2 (1 - 2\sqrt{2} M l_0)}} \frac{\partial \Theta}{\partial \psi}. \quad (59)$$

Then, using Eq. (59) in Eq. (55) and Eq. (54), our final expressions for the area and angular-diameter distance are

$$dA_s = \frac{\sin \Theta}{2l^2} \left| \frac{\partial \Theta}{\partial \psi} \right| \left(1 - \sin^2 \psi \frac{l^2(1 - 2\sqrt{2}Ml)}{l_0^2(1 - 2\sqrt{2}Ml_0)} \right)^{\frac{1}{2}} d\psi d\gamma \quad (60)$$

and

$$D_A^2 = \frac{\sin \Theta}{2l^2 \sin \psi} \left| \frac{\partial \Theta}{\partial \psi} \right| \left(1 - \sin^2 \psi \frac{l^2(1 - 2\sqrt{2}Ml)}{l_0^2(1 - 2\sqrt{2}Ml_0)} \right)^{\frac{1}{2}}. \quad (61)$$

It should be noted that the only place where D_A vanishes is at $\sin \Theta = 0$, namely, along the optical axis. The factor $\partial \Theta / \partial \psi$ does not vanish anywhere and diverges at $l = l_p$ at the same rate as the factor $(1 - \sin^2 \psi \frac{l^2(1 - 2\sqrt{2}Ml)}{l_0^2(1 - 2\sqrt{2}Ml_0)})^{1/2}$ approaches zero (See Eq. (A1) in the appendix).

From Eq. (52) and Eq. (61), we see that the square of the angular-diameter distance is proportional to the Jacobian of the lens mapping. Because the angular diameter distance appears in the denominator of the apparent brightness, S , (see Eq. (13)), a point source lying on the caustic will be infinitely magnified in the geometrical optics limit. In addition, our expression for the angular-diameter distance substituted into Eq. (14) gives the relative magnifications for two lensed images.

If one observes two or more images in the directions $\{\psi_i\}$, Eq. (37) can be used to define time of arrivals, $\{t_i\}$. The subtraction of two such times defines a coordinate time delay, which can be converted into a proper time delay along an observer's world line.

Among other possible candidates to useful lensing observables, which we have not concerned ourselves with, preliminary calculations suggest that the distortion of the images of small sources could be suitable for the application of the exact formalism as developed in this particular section.

C. Observables and the Parameter l

In each of our calculations of observable quantities, the non-measurable inverse radial parameter, l , plays an essential role. As mentioned earlier, this parameter should be eliminated in terms of observable quantities, perhaps a physical distance scale.

The most direct possibility is the angular-diameter distance expression given by Eq. (61). We can see that it is not a simple matter to invert the angular-diameter distance in order to infer l in terms of observables and lens properties. Nevertheless, Eq. (61) is an implicit relationship between l and the observable D_A , and can be solved numerically in local patches.

A second observational way to estimate the value of l , is via Eq. (14). We have indicated that the ratio of the brightness of a source in two images yields an implicit equation for the source position l through the distance relationship, Eq. (61). Hence, the parameter l may be replaced by μ_{12} in all calculations.

[An alternative approach, perhaps of only academic interest, to the inverse radial distance l can be obtained from the redshift of the source in closed form. If we assume that the source and observer are at rest then $v_s^a = |g_{00}(x^a(\alpha_1, \alpha_2, x^1))|^{-1/2}(1, 0, 0, 0)$ and $v_0^a = |g_{00}(x_0^a)|^{-1/2}(1, 0, 0, 0)$, thus $g_{ab}v^a \dot{X}^b = |g_{00}|^{-1/2}g_{00}\dot{t}$ for both ω_s and ω_0 . Using the metric, we also have $g_{00}\dot{t} = 1/\sqrt{2}$ at both locations. Thus

$$\frac{\omega_s}{\omega_0} = \left(\frac{1 - 2\sqrt{2}Ml_0}{1 - 2\sqrt{2}Ml} \right)^{\frac{1}{2}}, \quad (62)$$

which is of course the standard gravitational redshift [8] for Schwarzschild spacetime, expressed in our notation. Thus l can be obtained as a function of the ratio of the observed and source frequencies (and lens mass M and observers position l_0) by

$$l = \frac{1}{2\sqrt{2}M} \left(1 - \frac{\omega_s^2}{\omega_0^2} (1 - 2\sqrt{2}Ml_0) \right). \quad (63)$$

D. Image and Lens Properties

If one is interested in learning properties of an unseen lens, then the lens equation Eq. (7) with Eq. (61) can be used in conjunction with knowledge of the image properties, especially the brightness of the two main images. Because the brightness S of the images is proportional to an inverse power of the angular-diameter distance D_A , (via Eq. (13)), the brighter images will be observed when the source is located near a caustic.

We can see that the angular-diameter distance vanishes at locations where either M_1^a or M_2^a vanish. At such locations, neighboring rays meet. We have that M_1^a vanishes at $\Theta = 0$, which means that neighboring rays with the same value of ψ meet along the optical axis. (In fact, all rays with that value of ψ cross there, although only neighboring ones contribute to the intensity of the image.) This means that an observer in line with the lens and a source would observe an extremely bright perfect ring centered on the optical axis, at an angle ψ that makes the function Θ vanish (the Einstein ring). From Fig. 3, it can also be seen that as the image angle approaches zero, the source angle, Θ , tends to infinity. This means that there will be an infinite number of Einstein rings, appearing in principle for a black hole lens, at smaller and smaller angles – one each time that Θ passes through an integer number of turns, i.e., $\Theta = 2n\pi$. However, $\partial\Theta/\partial\psi$ goes to infinity as well, which means that the additional Einstein rings get much dimmer as the image angle goes to zero.

If the source is not on the optical axis, two main images will form, one on each side of the lens, because Θ is an odd function of ψ . One image will form at large positive angle ψ , whereas the opposite image forms at small negative angle ψ , and they will have different brightness. The images at smaller angles are dimmer, because Θ diverges steeply at the $3M$ radius, i.e., $\partial\Theta/\partial\psi$ is large. As with the Einstein rings, in addition to the two images there will be an infinite number of other images, dimmer and at smaller angles for a black hole lens.

On the other hand $M_2^a = 0$ would vanish at locations such that $\partial\Theta/\partial\psi = 0$, where neighboring rays with the same value of γ meet. These lie on a plane containing the source, the observer and the lens. It can be seen from Fig. 3 that in the case of a black hole, where the mass is contained within the $3M$ radius, Θ is a monotonically increasing function of ψ , and thus there are no points where $\partial\Theta/\partial\psi$ vanishes. Thus, we are not concerned with these caustics. These caustics do form in the case where a spherical lens is modeled as a uniform dust sphere with radius larger than $3M$, and lie on the light rays that travel through the mass, assuming the mass is transparent.

We can now use the lens equation Eq. (7) to infer properties of an unseen dark matter or blackhole lens. Using Eq. (61) with Eq. (7), and labeling the optical axis as the z -axis, we have:

$$\theta = \Theta(l(M, \text{observables}), l_0, \psi) \quad (64)$$

where the *observables* might be the brightness and luminosity or (see below) the frequency ratios. For example, if an Einstein ring is observed at an angle ψ_1 , then we know both the image angle and the source's angular location, i.e., $\theta = 0$. Then Eq. (64) yields a relationship between the mass M and the inverse radial location l_0 of the unseen blackhole. This is not particularly useful, but if another Einstein ring is observed at an angle ψ_2 , then we have two equations,

$$0 = \Theta(l(M, \text{observables}), l_0, \psi_1) \quad (65)$$

$$0 = \Theta(l(M, \text{observables}), l_0, \psi_2) \quad (66)$$

for the two unknowns l_0 and M and we can infer both the location and mass of the blackhole from lensing observables. Notice that this exact method is necessary in order to treat multiple Einstein rings, since the standard weak-field lens equation yields only one Einstein ring. This fact has been observed in [5], and is emphasized in the following section, where we compare the standard lens equation with the exact approach. As for the applicability of the method, we refer the reader to [5], where magnifications and angular locations of multiple Einstein rings have been calculated for the case of the Milky Way's galactic blackhole.

IV. COMPARISONS OF THE EXACT AND THIN-LENS EQUATIONS

The available approaches to lensing build on the view that the lens is a perturbation on a given background. The backgrounds are normally taken as Minkowski spacetime or as a Friedmann model. For definiteness, here we restrict to a Minkowski background. Given the flat background, one can define the locations of the source plane and the lens plane, at distances D_s and D_d from the observer, respectively, as shown in Fig. 4. One can further define the angular location of the source on the lens plane β , and the angular location of the image on the lens plane ψ . One can also define D_{ds} , the distance between the lens plane and the source plane. The thin-lens approximation consists in considering the bending to take place at the lens plane, the light rays being otherwise straight lines. The straight

lightrays bend through a bending angle α . The thin-lens approximation is justified from the point of view that the distances involved are much larger than the extent of the gravitational field, and is normally accompanied by the assumption that the image angles are small.

In our comparisons, we consider two different approximate lens equations available. One is the standard weak-field lens equation [1], and the other is a strong-field thin lens equation obtained recently in [5]. These two lens equations differ essentially in the calculation of the bending angle at the lens plane.

The standard weak-field approximation calculates the bending angle via linearized Schwarzschild. This results in small bending angles, which justifies a further assumption that the source angle β is small. Thus the standard weak-field thin-lens equation is also a small-angle approximation, where $\tan x = \sin x = x$ for $x = \beta, \psi, \alpha$. The weak-field thin-lens equation for a linearized Schwarzschild lens is

$$\beta = \psi - \frac{4MD_{ds}}{D_d D_s \psi} \quad (67)$$

In our current notation, the angular location of the source from the optical axis is θ . We need to transform β into a function of θ in order to make a comparison with the exact lens equation. In this approximation, however, since all angles are small, then from Fig. 4. $\beta/\theta = D_{ds}/D_s$, thus

$$\beta = \frac{D_{ds}}{D_s} \theta. \quad (68)$$

We also need to express the distances in terms of the coordinates (l, θ) . In this case, because of the small-angle assumption, we have

$$D_d = \frac{1}{\sqrt{2}l_0}, \quad (69)$$

$$D_{ds} = \frac{1}{\sqrt{2}l}, \quad (70)$$

$$D_s = \frac{1}{\sqrt{2}l_0} + \frac{1}{\sqrt{2}l}. \quad (71)$$

Using Eq. (68) and Eq. (71) with Eq. (67), we obtain the weak-field thin-lens equation in our current notation:

$$\theta = \frac{l+l_0}{l} \left(\psi - \frac{4\sqrt{2}M}{\psi} \frac{l_0^2}{(l_0+l)} \right). \quad (72)$$

This is to be compared with the lens equation, Eq. (41), in the case $\theta_0 = \pi$. We can see that $\theta = \Theta$ reduces to Eq. (72) in the following manner. We assume that the dimensionless quantities $M l \equiv \epsilon$ and $M l_0 \leq \epsilon$ are small and make a Taylor series expansion of $\Theta \equiv \pi - \Delta$ in terms of ϵ :

$$\Theta(l, l_0, \psi) \approx \pi - \Delta(\epsilon=0, b^*, l_0, l) + \epsilon \left[\frac{\partial \Delta}{\partial \epsilon} \right]_{\epsilon=0} + \frac{\epsilon^2}{2} \left[\frac{\partial^2 \Delta}{\partial \epsilon^2} \right]_{\epsilon=0} \equiv \pi - \Delta_0 - \Delta_1 - \Delta_2. \quad (73)$$

One can show that the zeroth and first order terms in the Taylor series expansion, Eq. (73), yield Eq. (72). We can further derive the next term in the expansion. The second-order correction, Δ_2 , is obtained to lowest order in ψ by

$$\Delta_2 = \frac{15 \pi (M)^2}{4 D_d^2 \psi^2}, \quad (74)$$

which corrects Eq. (72) as

$$\theta = \frac{l+l_0}{l} \left(\psi - \frac{4\sqrt{2}M}{\psi} \frac{l_0^2}{(l_0+l)} - \frac{15\pi M^2}{2\psi^2} \frac{l_0^3}{l_0+l} \right). \quad (75)$$

Equation (75) translates into a second-order accurate version of the standard weak-field thin-lens equation in the form

$$\beta = \psi - \frac{4M}{D_d} \frac{D_{ds}}{D_s \psi} - \frac{15\pi}{4} \frac{M^2}{D_s} \frac{D_{ds}}{D_d^2 \psi^2}. \quad (76)$$

The strong-field thin-lens equation [5] calculates the bending angle in exact Schwarzschild, for a lightray that deviates a total angle $\hat{\alpha}$ between the incoming direction from infinity and the outgoing direction to infinity. The bending angle $\hat{\alpha}$ is not necessarily small, and will increase to infinity as rays approach the $3M$ radius. Thus there is no basis for a small angle approximation of the angles β and $\hat{\alpha}$. The image angle ψ is still considered small in the spirit of the thin lens approximation, and the lens and source spheres are still considered planes. The strong-field thin-lens equation is obtained in the same manner as the weak-field thin-lens equation is, from trigonometry on the lens diagram, but uses the exact trigonometric functions of β and $\hat{\alpha}$, rather than substitute the sines and tangents by their angles. The strong-field thin-lens equation is Eq. (1) in [5], namely

$$\tan \beta = \tan \psi - \frac{D_{ds}}{D_s} [\tan \psi + \tan(\hat{\alpha} - \psi)] \quad (77)$$

with

$$\hat{\alpha} = 2 \int_0^{l_p} dl \frac{1}{\sqrt{l_p^2(1 - 2\sqrt{2}Ml_p) - l^2(1 - 2\sqrt{2}Ml)}} - \pi \quad (78)$$

Equation (78) is obtained from Eq. (10) in [5] by making the identifications $r \rightarrow \frac{1}{\sqrt{2}l}$ and $r_0 \rightarrow \frac{1}{\sqrt{2}l_p}$.

Since we are interested in comparing with the exact lens equations, we need to express $\beta, \hat{\alpha}$ and D_d, D_{ds}, D_s in terms of our coordinates. In the first place, since angles are not small then we have

$$\tan \beta = \frac{D_{ds}}{D_s} \tan \theta \quad (79)$$

and

$$D_d = \frac{1}{\sqrt{2}l_0}, \quad (80)$$

$$D_{ds} = \frac{\cos \theta}{\sqrt{2}l}, \quad (81)$$

$$D_s = \frac{1}{\sqrt{2}l_0} + \frac{\cos \theta}{\sqrt{2}l}. \quad (82)$$

Additionally, from Eq. (39), it is straightforward to see that

$$\hat{\alpha} = -\Theta(l, l_0, \psi)|_{l=l_0=0}. \quad (83)$$

With Eqs. (79), (82) and (83), Eq. (77) becomes

$$\sin \theta - \cos \theta \tan (\Theta|_{l=l_0=0} + \psi) - \frac{l}{l_0} \tan \psi = 0 \quad (84)$$

which can be manipulated to yield

$$\theta = \psi + \Theta|_{l=l_0=0} + \arcsin \left(\frac{l}{l_0} \tan \psi \cos (\Theta|_{l=l_0=0} + \psi) \right). \quad (85)$$

Equation (85) is the strong-field thin-lens equation and is to be compared with the exact lens equation $\theta = \Theta$.

We can now plot our three lens equations and see how well they compare. We choose $l = l_0 = \frac{0.001}{3\sqrt{2}M}$, which corresponds to a distance of $3000M$ from the lens for both the source and the observer and set $M = 1$ for the plots. We look at small image angles, which corresponds to approaches close to the $3M$ radius. The plots are only made for positive image angles. Since the lens equations are odd functions of ψ , the plot for negative ψ is a reflection through the origin and can be inferred from Fig. 5.

We can see in Fig. 5 that the weak-field equation (72) deviates significantly from the exact equation both at first-order accuracy and at second-order accuracy, Eq. (75), as expected. The main reason for the second-order correction not to behave significantly better than the first-order accurate equation is that none of the perturbative terms blow up at the $3M$ radius, whereas the exact equation does. The perturbative approach is clearly inappropriate for the strong-field regime of a black hole.

However, the strong-field thin-lens equation Eq. (85) agrees remarkably well with the exact equation. In our example, we have an error of less than one part in a thousand, for source angles as large as about six turns around the lens. Notice that 3000M corresponds to only 4500km for a stellar black hole, nowhere near the galactic distances expected for observed lensing. The thin-lens approximation appears to do very well at these relatively small distances and will almost certainly do much better at galactic and extragalactic distances.

One reason why Eq. (77) performs so accurately is the tremendous distances involved. The other reason is, however, that the exact bending angle through infinite distances is used. Most of the effort in using Eq. (77) is involved in evaluating the integral expression of the bending angle and its derivatives, which involves the same amount of work required to do the exact lens equation, as shown by our comparison above, Eq. (85). Thus the efficiency in terms of computational cost is not very high. Considering that the price is to break with the covariance and non-linearity of general relativity, one might even find it justified to give up the thin-lens approximation in favor of the exact lens equation in the case of strong, spherically symmetric gravitational fields.

A valuable, second comparison is the time delays predicted by the exact and thin lens approaches. Time delays are now a very important tool in astrophysics and provide a means to estimate the Hubble constant independent of previous methods. We will define a time delay as the observer's elapsed proper time between the arrival of a signal along two distinct paths connecting the source and observer.

When the weak field, thin lens approach is applied to the Schwarzschild lens, generically, two paths connect any source and observer, one passing on each side of the lens. This results from the lens equation, Eq. (41), admitting two values for ψ for given values of β , M , and distance parameters. The only exception is when the source, lens, and observer lie along the same radial line. In this case an Einstein ring is formed, and the time delay is zero by symmetry. The time along each of the two paths is found by integrating

$$t^{tl} = \int dl \left(1 + \frac{2M}{r(l)} \right) \quad (86)$$

along the trajectory determined by the weak field thin lens equation. In the integration, l is the Euclidean length along the thin lens trajectory, and $r(l)$ represents the Euclidean distance from the origin to a point along the path. The thin lens time delay, Δt^{tl} , is the difference between the two times.

In the exact approach, there are an infinite number of pairs of geodesics connecting the source and observer because geodesics may wrap around the black hole many times. However, there are only two geodesics which do not circle the lens; these geodesics, which can be distinguished from the others, are the analogue of the thin lens paths. The exact time delay, Δt^e , is numerically computed from Eq. (37) by taking the difference between the times elapsed along these two geodesics.

Figure 6 shows the exact and weak field thin lens time delays given the same observer and source location. The time delays are given in geometrical units, and the β -axis is given in radians. In this calculation, the observer is located at $l_0 = \frac{0.001}{3\sqrt{2}M}$, or a distance of 3000M, and $\theta = 0$. Sources are located in a fixed source plane whose distance from the lens along the optical axis is also 3000M, as in the standard thin lens picture. The time delays are plotted against β , the source angle used in the thin lens equation.

At $\beta = 0$, Einstein rings are formed, and each time delay will be zero. As β increases, the thin lens time delay slightly overestimates the true value, as seen in the figure. This represents the general behavior for a single Schwarzschild lens at larger distances, but the effect remains fairly small. For a single Schwarzschild lens with a mass of 2.5×10^{12} solar masses, at redshift $z = 0.5$ with a source at $z = 1.0$ and $\beta = 0.5''$, the overall exact time delay is close to 400 days, while the error introduced by using the thin lens approximation is less than an hour.

The time delays computed in the thin lens approximation remain quite accurate in our comparisons because the light rays do not feel a strong gravitational field along their trajectories. This is why the thin lens fails so dramatically in the observation angle calculation presented in Fig. 5, while fairing well in the time delay comparisons.

Despite these results, it may be incorrect to assume that the thin lens method of computing time delays remains accurate for more complex lensing situations. The first of two possible problems is that the weak field assumption may not always be valid along an observed null geodesic. It may be very difficult to identify lens candidates whose null geodesics have undergone interactions with strong gravitational fields, however, if such candidates are found, the thin lens methodology will most likely prove quite inaccurate. Secondly, it is not known how accurate the time delay computations will be in more complex lensing scenarios, especially when the geodesics are bent in multiple lens planes.

V. DISCUSSION

To date, virtually all applications of gravitational lensing have utilized the thin lens approximation. Although the thin lens method has proven a quick and useful tool, it can not be applied to high curvature regions. The failure

of the next order correction to the thin lens equation, Eq. (76), as contrasted with the ability of the strong field thin lens equation [5], to capture the divergence of the observation angle Θ in Fig. 5 emphasizes the point that some combination of exact methods are required in such situations.

We are putting forward the idea that exact gravitational lensing may not be just a purely academic exercise. In fact, we see that virtually all of the observationally relevant quantities can be determined analytically in the exact method with relative ease for a Schwarzschild spacetime. In this paper, we found straight-forward expressions for the time delays, observation angles, angular-diameter distances, and relative magnifications. These analytic expressions can be used in comparison calculations or model building with great computational prowess, power, or time.

The limited testing of the thin-lens methodology in this paper has indicated that there are regions (possibly not yet observable) where the thin-lens method fails in a highly symmetric case. In ongoing calculations, we are exploring the accuracy of the thin lens approximation under a broad range of conditions, including those in which there may be stronger fields or multiple lenses present. These calculations will require a combination of analytic and numerical results. A precedent for this kind of work was set by Rauch and Blandford [10], who studied the null geodesic equations (or the exact lens equations in our terminology) for Kerr spacetime and found the caustics of the lightcone. Our sense from these comparisons is that the error in the thin lens method for the time delays and observation angles will, in some cases, be appreciable [9].

We are also interested in an issue regarding the comparison of the sizes of two different types of corrections to the thin lens equations. On one hand, within the framework of the thin lens methodology, there are corrections due to the structure of the mass distribution (of the lens) over that of the monopole moment. On the other hand, even for the monopole case, there are differences between the predictions of the exact lens equations and the thin lens equations. The issue is whether the sizes of these corrections are comparable.

ACKNOWLEDGMENTS

This research has been supported by the NSF under grants No. PHY-9803301, PHY-9722049 and PHY-9205109.

APPENDIX A:

An alternative derivation of Eq. (59), by direct computation, is obtained by taking the derivative of Θ and of t with respect to ψ . Because both t and Θ depend on ψ only through the point of closest approach l_p , we actually need to calculate $\partial t/\partial l_p$ and $\partial \Theta/\partial l_p$. We have

$$\begin{aligned} \frac{\partial \Theta}{\partial l_p} = & + \int_{l_0}^{l_p} \frac{l_p(1 - 3\sqrt{2}Ml_p)dl'}{\left(l_p^2(1 - 2\sqrt{2}Ml_p) - l'^2(1 - 2\sqrt{2}Ml')\right)^{3/2}} \\ & + \lim_{\epsilon \rightarrow 0} \left\{ 2 \int_{l_0}^{l_p - \epsilon} \frac{l_p(1 - 3\sqrt{2}Ml_p)dl'}{\left(l_p^2(1 - 2\sqrt{2}Ml_p) - l'^2(1 - 2\sqrt{2}Ml')\right)^{3/2}} \right. \\ & \left. - \frac{2}{\sqrt{l_p^2(1 - 2\sqrt{2}Ml_p) - l'^2(1 - 2\sqrt{2}Ml')}} \right|_{l'=l_p - \epsilon} \Bigg\} \end{aligned} \quad (A1)$$

On the other hand we also have

$$\begin{aligned} \frac{\partial t}{\partial l_p} = & - \frac{1}{\sqrt{2}} \int_{l_0}^{l_p} \frac{l_p(1 - 3\sqrt{2}Ml_p)}{\sqrt{l_p^2(1 - 2\sqrt{2}Ml_p)}} \left(l_p^2(1 - 2\sqrt{2}Ml_p) - l'^2(1 - 2\sqrt{2}Ml')\right)^{3/2} dl' \\ & - \lim_{\epsilon \rightarrow 0} \left\{ \frac{2}{\sqrt{2}} \int_{l_0}^{l_p - \epsilon} \frac{l_p(1 - 3\sqrt{2}Ml_p)}{\sqrt{l_p^2(1 - 2\sqrt{2}Ml_p)}} \left(l_p^2(1 - 2\sqrt{2}Ml_p) - l'^2(1 - 2\sqrt{2}Ml')\right)^{3/2} dl' \right. \\ & \left. + \frac{2}{\sqrt{2}l'^2(1 - 2\sqrt{2}Ml')} \sqrt{\frac{l_p^2(1 - 2\sqrt{2}Ml_p)}{l_p^2(1 - 2\sqrt{2}Ml_p) - l'^2(1 - 2\sqrt{2}Ml')}}} \right|_{l'=l_p - \epsilon} \Bigg\} \end{aligned} \quad (A2)$$

Clearly the last term inside the braces under the limit sign in (A2) can be substituted with

$$\frac{2}{\sqrt{2l_p^2(1 - 2\sqrt{2}Ml_p)}} \frac{1}{\sqrt{l_p^2(1 - 2\sqrt{2}Ml_p) - l'^2(1 - 2\sqrt{2}Ml')}} \Big|_{l'=l_p-\epsilon} \quad (\text{A3})$$

since the difference vanishes in the limit $\epsilon \rightarrow 0$. We can thus see that $\partial t/\partial l_p$ and $\partial \Theta/\partial l_p$ are proportional to each other via

$$\frac{\partial t}{\partial l_p} = - \frac{1}{\sqrt{2l_p^2(1 - 2\sqrt{2}Ml_p)}} \frac{\partial \Theta}{\partial l_p}. \quad (\text{A4})$$

This implies Eq.(59).

-
- [1] P. Schneider, J. Ehlers, and E. E. Falco. *Gravitational Lenses*, (Springer-Verlag, New York, Berlin, Heidelberg, 1992).
 - [2] R. Blanford and N. Gehrels. *Physics Today*. **52**, vol. 6, 40, (1999)
 - [3] S. Frittelli and E. T. Newman. *Phys. Rev. D* **59**, 124001, (1999).
 - [4] J. Ehlers, S. Frittelli, and E. T. Newman. “Gravitational Lensing From a Space-Time Perspective”, to appear in the *Festschrift for John Stachel*, Ed. J. Renn, Kluwer Academic Publishers, (2000)
 - [5] K. S. Virbhadra and G. F. R. Ellis, astro-ph/9904193.
 - [6] E. Schrödinger. *Expanding Universes*. Cambridge University Press, NY, 1956.
 - [7] T. P. Kling and E. T. Newman. *Phys. Rev. D* **59**, 124002, (1999).
 - [8] Robert Wald. *General Relativity*. Univ. of Chicago Press, Chicago, 1984.
 - [9] T. P. Kling, E. T. Newman, A. Perez. *in preparation*
 - [10] K. P. Rauch and R. Blandford. *ApJ* **421**, 46, (1994).

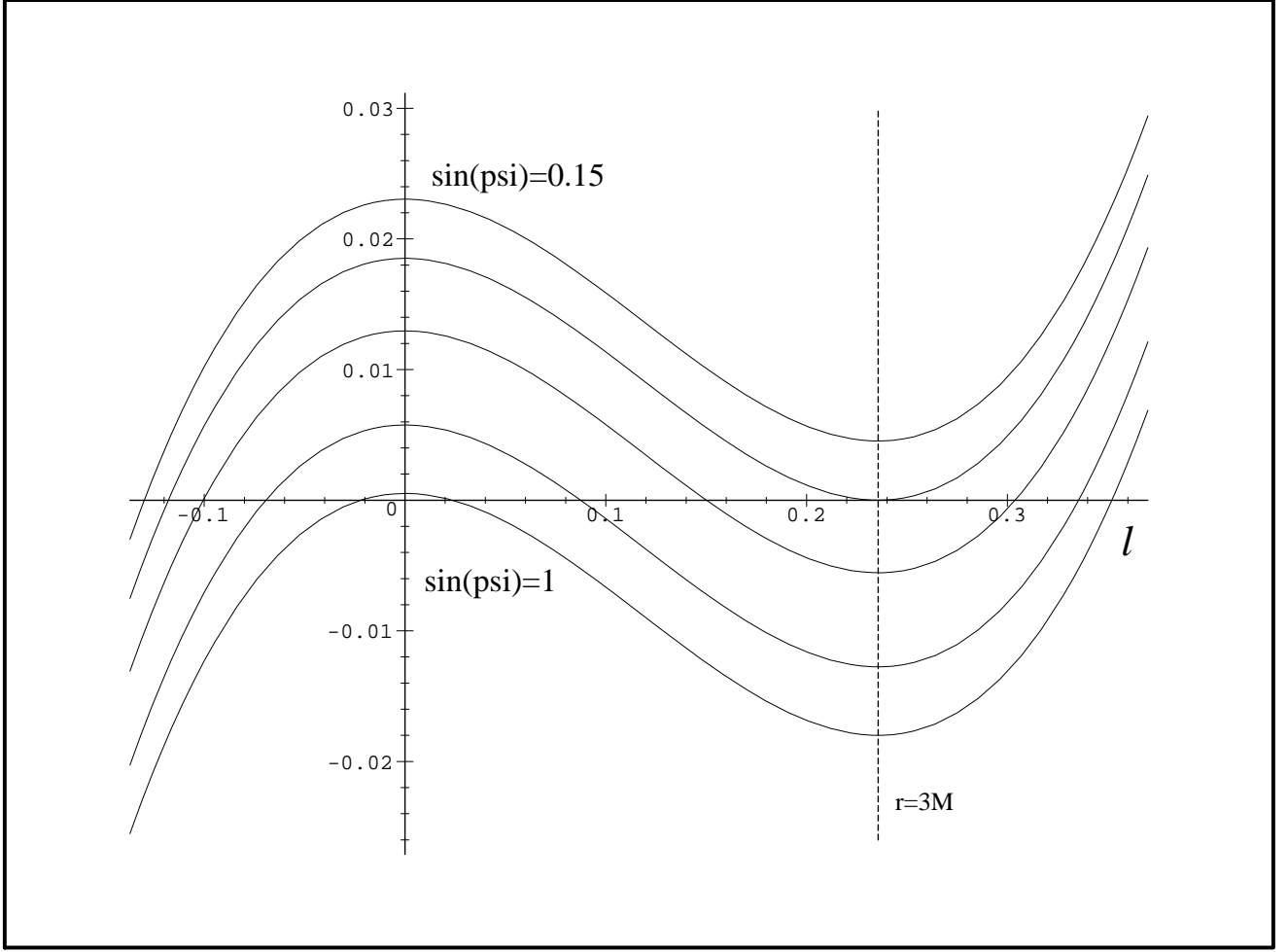


FIG. 1. A plot of the function $F := 1 - \sin^2 \psi l^2 f / (l_0^2 f(l_0))$ as a function of l for a sequence of values of $\sin \psi$ including $\sin \psi = 1$. Notice that for large enough ψ there are always two positive roots, while for small enough ψ there are no positive roots. There is a critical value of ψ for which there is only one positive root. The allowed values of l are such that F is positive, and such that $l = 0$ is included. Thus, allowed values of l range from 0 to the smallest positive root. The smallest positive root is denoted l_p and represents the point of closest approach to the lens. Notice that l_p increases with decreasing ψ , reaching the critical value $(3\sqrt{2}M)^{-1}$ at the critical value of ψ .

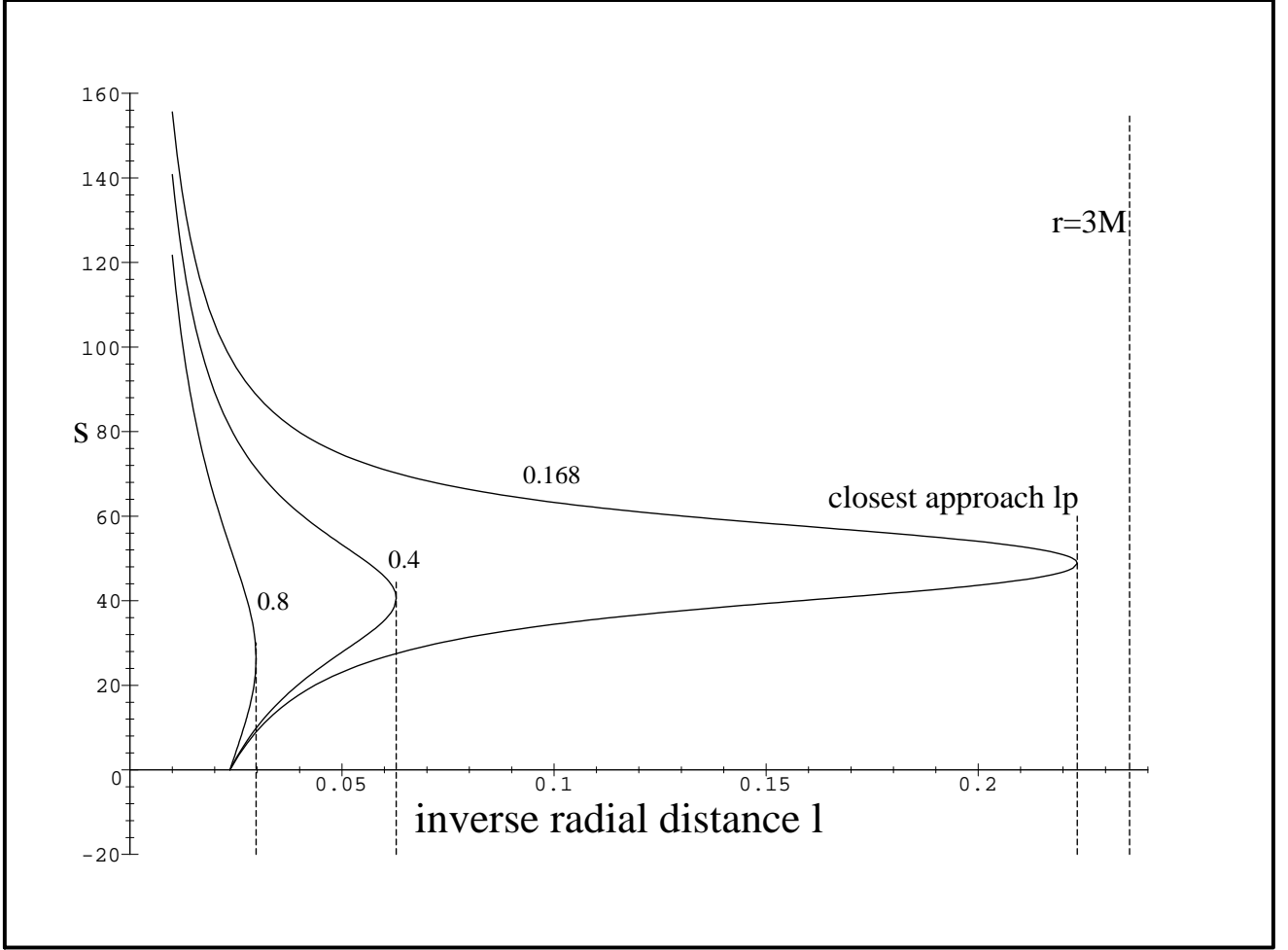


FIG. 2. The affine length s as a function of the inverse radial distance l for three null geodesics that reach the observer from positive directions ψ . We have set the observer at $l_0 = 0.1/3\sqrt{2}M$, the mass is $M = 1$, and the constant C has the value 1 for the three null geodesics. The three null geodesics are labeled according to the value of $\sin \psi$. Smaller positive angles ψ reach closer to the $3M$ radius at their respective point of closest approach l_p . For this observer the smallest image angle is at $\sin \psi = 0.16733201$, at which the closest approach reaches the $3M$ radius. We see that the affine length is double valued as a function of the inverse radial distance.

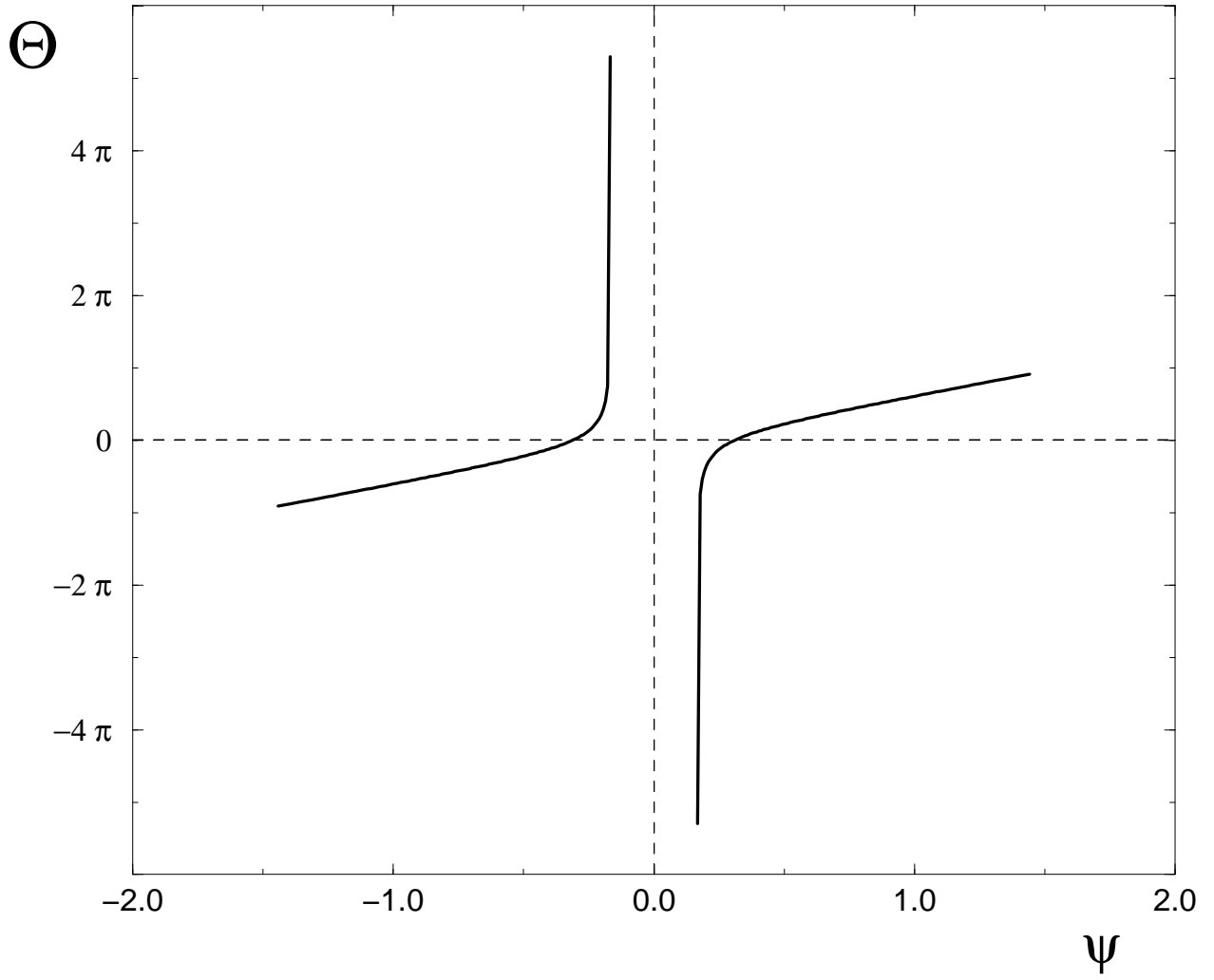


FIG. 3. The function $\Theta(l, l_0, \psi)$ at fixed values of l and l_0 . The angles on both axis are measured in radians. We have chosen $l = l_0 = 0.1/3\sqrt{2}M$ with $M = 1$ for this figure. This plot represents the exact lens equation in the case that the z -axis is chosen as the optical axis, defined by the observer and the lens.

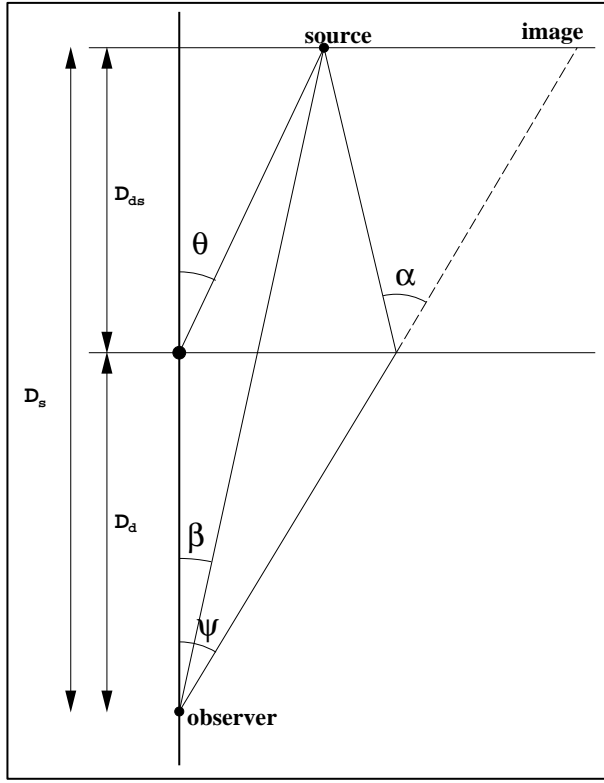


FIG. 4. The lens diagram for the case in which a background is available. The quantities $\beta, \psi, \alpha, D_s, D_d, D_{ds}$ represent the Euclidean angles and angular diameter distances in the background space-time shown. The quantity θ represents the angular position of the source with respect to the z -axis in Schwarzschild coordinates if Schwarzschild spacetime is thought of as superimposed on a flat background.

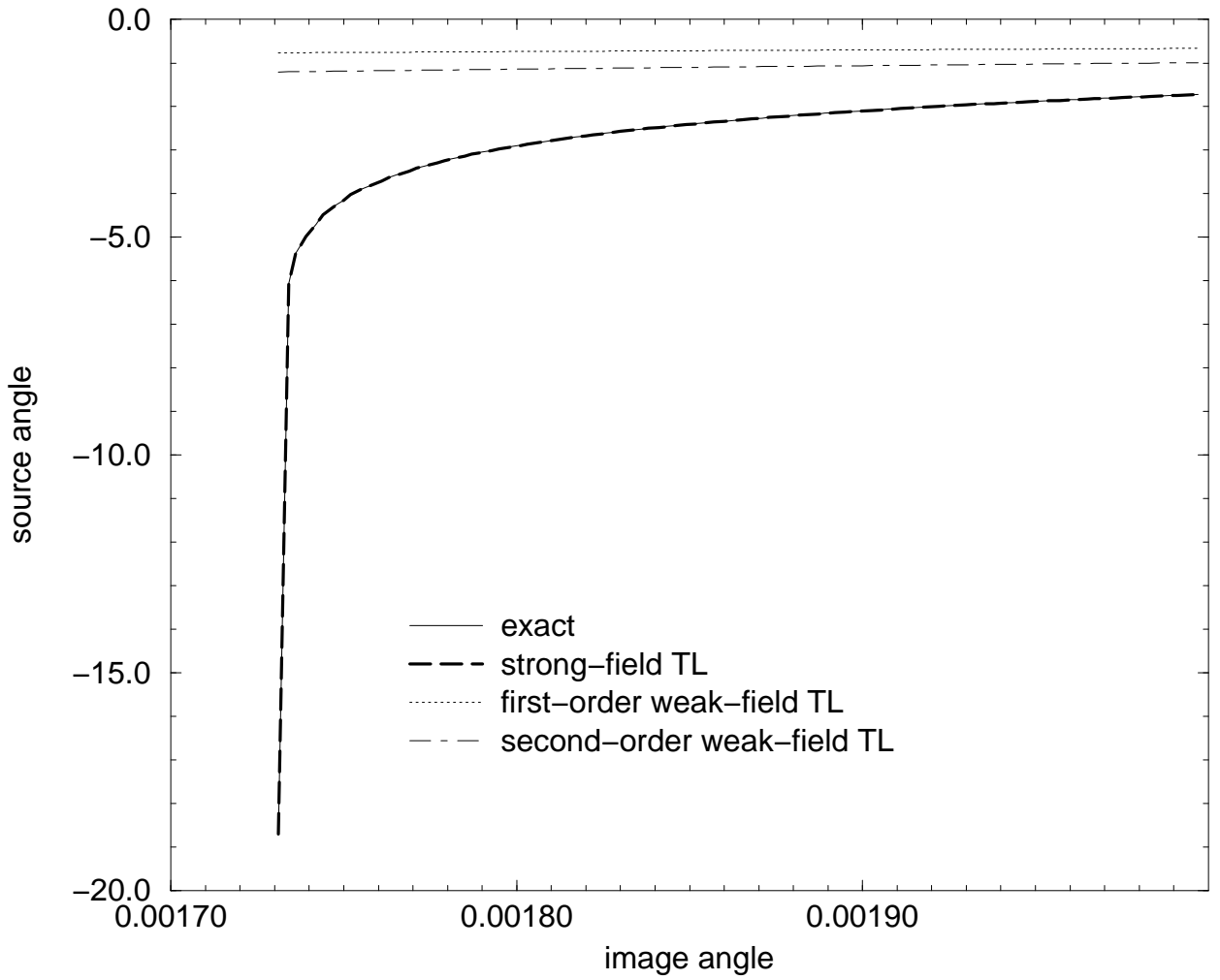


FIG. 5. A comparison of the available lens equations in the literature. The vertical axis represents the angular position of the source in radians. The horizontal axis represents the observed angle ψ **in radians**. We have taken the observer and the source at the same distance from the lens, $r = 3000M$. Here TL:=thin lens. The standard thin-lens-weak-field approximation is clearly inaccurate, whereas the second-order correction to it does not do significantly better. However, the thin-lens-strong-field approximation appears to be remarkably good, since at this resolution it agrees with the exact lens equation up to one part in a thousand.

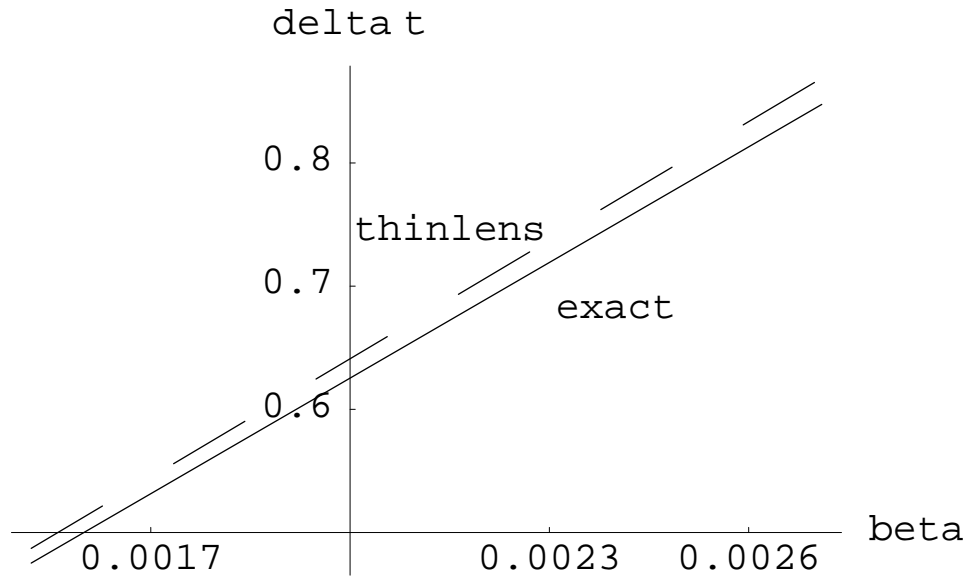


FIG. 6. A plot of the exact and thin lens time delays against the source angle β in radians. Here, the source and observer are roughly $3000M$ from the lens.